

Claims

What is claimed is:

- 5
1. A method of managing a HTTP cache in a web server, comprising:
- receiving a HTTP request;
- determining whether response data associated with the HTTP request
- is in the HTTP cache;
- when it is determined that response data associated with the HTTP
- 10 request is in the HTTP cache, obtaining an advisory state associated with the
- HTTP request from the HTTP cache, the advisory state when in a first state
- indicating that it is necessary to consult with the web server to determine
- whether the response data can be transmitted and when in a second state
- 15 indicating that the response data can be transmitted without consulting the
- web server; and
- transmitting the response data in accordance with the advisory state
- associated with the HTTP request.
- 20
2. The method as recited in claim 1, wherein transmitting the response
- data in accordance with the advisory state associated with the HTTP request
- comprises:
- when the advisory state is in the first state, sending an advisory request
- 25 to a HTTP daemon, the advisory request requesting an advise state from the
- HTTP daemon indicating an action to be taken with the response data, and
- receiving the advise state from the HTTP daemon.
- 30

3. The method as recited in claim 2, wherein transmitting the response data in accordance with the advisory state associated with the HTTP request when the advisory state is in the first state further comprises:

transmitting the response data without modifying the response data in the HTTP cache when the advise state is in a first state.

4. The method as recited in claim 2, wherein transmitting the response data in accordance with the advisory state associated with the HTTP request when the advisory state is in the first state further comprises:

modifying the response data stored in the HTTP cache as specified by the advise state.

5. The method as recited in claim 4, wherein modifying the response data comprises:

removing at least one of the response data and the advisory state from the HTTP cache when the advise state is in a second state.

6. The method as recited in claim 4, wherein modifying the response data comprises:

receiving second response data from the HTTP daemon; and
performing at least one of replacing the response data in the HTTP cache with the second response data and replacing the advisory state in the HTTP cache with a second advisory state when the advise state is in a third state.

7. The method as recited in claim 2, wherein transmitting the response data in accordance with the advisory state associated with the HTTP request when the advisory state is in the first state further comprises:

receiving second response data from the HTTP daemon; and
transmitting the second response data when the advise state is in a
fourth state without transmitting the response data in the HTTP cache and
without storing the second response data in the HTTP cache.

5

8. In a web server, a method of managing an in-kernel cache, comprising:
receiving a request;
10 determining whether response data associated with the request is in the
cache; and

when it is determined that the response data associated with the request
is not in the cache, sending the request to a process and receiving response
data associated with the request and one or more associated caching attributes
15 from the process, the caching attributes including a cache state indicating that
the response is to be stored in the cache when in a first state and the cache
state indicating that the response is not to be stored in the cache when in a
second state.

20 9. The method as recited in claim 8, wherein the in-kernel cache is an in-
kernel HTTP cache, the request is a HTTP request, and the process is a HTTP
daemon

10. The method as recited in claim 9, wherein the caching attributes
25 further include an advisory state associated with the HTTP request, the
advisory state when in a first state indicating that it is necessary to obtain
permission from the HTTP daemon in order to transmit the response data and
when in a second state indicating that the response data can be transmitted
without obtaining permission from the HTTP daemon.

30

11. The method as recited in claim 10, wherein the advisory state when in
the first state indicates that it is necessary to obtain permission from the HTTP

daemon to transmit response data that is stored in the HTTP cache and when in the second state indicates that the response data stored in the HTTP cache can be transmitted without consulting the HTTP daemon.

- 5 12. The method as recited in claim 10, further comprising:
 storing the advisory state and the response in the HTTP cache when
 the cache state is in the first state.

- 10 13. The method as recited in claim 9, wherein the caching attributes
 further include an identifier associated with the response to enable the
 response to be associated with multiple HTTP requests in the HTTP cache.

- 15 14. The method as recited in claim 13, further comprising;
 storing the identifier and the response in the HTTP cache when the
 cache state is in the first state.

- 20 15. A method of managing a HTTP cache in a web server, comprising:
 creating a cache request specifying a modification to be made to
 information stored in the HTTP cache;
 sending the cache request from a HTTP daemon to a cache manager;
 and
 modifying the information stored in the HTTP cache in accordance
25 with the cache request received by the cache manager from the HTTP daemon.

- 30 16. The method as recited in claim 15, wherein the HTTP cache is in a
 kernel of the web server.

17. The method as recited in claim 15, wherein the cache manager is in a
 kernel of the web server.

18. The method as recited in claim 15, wherein the cache request specifies an object to flush from the HTTP cache, the object including a HTTP response.

19. The method as recited in claim 18, wherein the object further includes an advisory state, the advisory state when in a first state indicating that it is necessary to consult with the HTTP daemon to determine whether the response data can be transmitted and when in a second state indicating that the response data can be returned without consulting the HTTP daemon.

20. The method as recited in claim 19, wherein the object further includes an identifier associated with the response to enable the response to be associated with multiple HTTP requests in the HTTP cache.

21. A web server including a system for controlling an in-kernel HTTP cache, comprising:

a HTTP daemon adapted for providing response data and one or more cache control indicators, the cache control indicators adapted for at least one of managing information that is stored in the HTTP cache and controlling transmission of the response data; and

a cache manager adapted for receiving the response data and the cache control indicators from the HTTP daemon and at least one of modifying the information stored in the HTTP cache and controlling the transmission of the response data in accordance with the cache control indicators.

22. The web server as recited in claim 21, wherein the cache control indicators include at least one of 1) an advisory state indicating whether the cache manager must communicate with the HTTP daemon in order to determine whether response data stored in the HTTP cache can be transmitted to a client that has sent a HTTP request, 2) a cache advise state indicating at

least one of that response data in the HTTP cache can be transmitted to the client, that a first set of data in the HTTP cache is to be replaced with a second set of data, that data in the HTTP cache is to be flushed from the HTTP cache, and that the response data is temporary and therefore is to be transmitted to the client and not stored in the HTTP cache, and 3) a cache state indicating whether the response data is to be stored in the HTTP cache

23. The web server as recited in claim 21, wherein the HTTP daemon is further adapted for independently providing at least one of the response data associated with the HTTP request and one or more of the cache control indicators without a request from the cache manager.

24. A computer-readable medium for managing a HTTP cache in a web server, the computer-readable medium storing computer-readable instructions thereon, comprising:

- instructions for receiving a HTTP request;
- instructions for determining whether response data associated with the HTTP request is in the HTTP cache;
- instructions for when it is determined that response data associated with the HTTP request is in the HTTP cache, obtaining an advisory state associated with the HTTP request from the HTTP cache, the advisory state when in a first state indicating that it is necessary to consult with the web server to determine whether the response data can be transmitted and when in a second state indicating that the response data can be transmitted without consulting the web server; and
- instructions for transmitting the response data in accordance with the advisory state associated with the HTTP request.

add d17